Hierarchical nanomechanics of collagen microfibrils

Alfonso Gautieri[1,2], Simone Vesentini[2], Alberto Redealli[2] and Markus J. Buehler[1,3,4†]

[1] *Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave. Room 1-235A&B, Cambridge, MA, USA*

[2] *Biomechanics Group, Department of Bioengineering, Politecnico di Milano, Via Golgi 39, 20133 Milan, Italy*

[3] *Center for Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA, USA*

[4] *Center for Computational Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA, USA*

[†] *Corresponding author, electronic address:* mbuehler@MIT.EDU*, Phone: +1-617-452-23750, Fax: +1-617-324-4014*

Abstract: Collagen constitutes one third of the human proteome, providing mechanical stability, elasticity and strength to connective tissues. Collagen is also the dominating material in the extracellular matrix (ECM) and is thus crucial for cell differentiation, growth and pathology. However, fundamental questions remain with respect to the origin of the unique mechanical properties of collagenous tissues, and in particular its stiffness, extensibility and nonlinear mechanical response. By using x-ray diffraction data of a collagen fibril reported by Orgel *et al*. (Proceedings of the National Academy of Sciences USA, 2006) in combination with protein structure identification methods, here we present an experimentally validated model of the nanomechanics of a collagen microfibril that incorporates the full biochemical details of the amino acid sequence of the constituting molecules. We report the analysis of its mechanical properties under different levels of stress and solvent conditions, using a full-atomistic force field including explicit water solvent. Mechanical testing of hydrated collagen microfibrils yields a Young's modulus of ≈300 MPa at small and ≈1.2 GPa at larger deformation in excess of 10% strain, in excellent agreement with experimental data. Dehydrated, dry collagen microfibrils show a significantly increased Young's modulus of ≈1.8 to 2.25 GPa (or ≈6.75 times the modulus in the wet state) owing to a much tighter molecular packing, in good agreement with experimental measurements (where an increase of the modulus by ≈9 times was found). Our model demonstrates that the unique mechanical properties of collagen microfibrils can be explained based on their hierarchical structure, where deformation is mediated through mechanisms that operate at different hierarchical levels. Key mechanisms involve straightening of initially disordered and helically twisted molecules at small strains, followed by axial stretching of molecules, and eventual molecular uncoiling at extreme deformation. These mechanisms explain the striking difference of the modulus of collagen fibrils compared with single molecules, which is found in the range of 4.8±2 GPa or ≈10-20 times greater. These findings corroborate the notion that collagen tissue properties are highly scale dependent and nonlinear elastic, an issue that must be considered in the development of models that describe the interaction of cells with collagen in the extracellular matrix. A key impact the atomistic model of collagen microfibril mechanics reported here is that it enables the bottom-up elucidation of structure-property relationships in the broader class of collagen materials such as tendon or bone, including studies in the context of genetic disease where the incorporation of biochemical, genetic details in material models of connective tissue is essential.

Keywords: Collagen; connective tissue; tendon; bone; mechanical properties; deformation; molecular simulation; nanomechanics; mesoscale; materiomics

Introduction

Collagen molecules represent the most abundant protein building block in the human body, where they provide mechanical stability, elasticity and strength to connective tissues such as tendons, ligaments and bone, as well as the extracellular matrix (ECM) [1]. Since pioneering experimental works by Fraser, Hulmes, Hess, Orgel, Fratzl and others it is known that virtually all collagen-based tissues are organized into hierarchical structures, where the lowest hierarchical level consists of triple helical collagen molecules (Fig. 1) [2-8]. Collagen fibrils consist of high-aspect-ratio polypeptides, tropocollagen molecules, with a length of ≈300 nm and a diameter of about 1.5 nm, which are laterally staggered. This structure creates an observable periodicity known as the *D*-band whose length is D=67 nm. The collagen molecule's length is not a multiple of D, where in terms of D the collagen molecule measures 4.46 D. According to the Hodge-Petruska model [9]—a structural model of collagen fibrils proposed in 1964—molecules in a fibril are deposited side by side and parallel but staggered with respect to each other, where the molecular axes are also parallel to the fibril direction. A gap between two consecutive collagen molecules is known as the "gap region" and measures 0.54 D, or ≈36 nm [9]. Collagen fibrils have a diameter of 100–500 nm and a length up to the millimeter range, and are formed through the bundling of several microfibrils that each contain clusters of five collagen molecules [10]. At the next level of the hierarchy, multiple fibrils make up the collagen fiber, formed with the aid of cross-linking macromolecules such as proteoglycans. In bone, the organic collagen protein matrix alone is not sufficient alone to provide the stiffness required for this tissue (which has to carry considerable loads) and additional stiffening is provided by the inclusion of mineral hydroxyapatite crystals into collagen fibrils that emerge from the gap region [5, 7, 11].

Significant efforts have been put forth in recent years focused on characterizing the mechanical properties of collagen by using experimental, computational and theoretical approaches. A review of recent efforts in the understanding of the structure and mechanical properties of collagen-based tissues is nicely summarized in recent works [7, 12]. Earlier work was mostly focused on the macroscopic, overall mechanical properties of collagen fibers and related tissues, with several efforts that elucidated the mechanics of hierarchical structures at intermediate scales [7, 11, 13-15]. Other studies focused solely on the properties of individual tropocollagen molecules without linking to the larger scale materials response [16-19]. Several recent reports focused on the mechanical properties of individual collagen fibrils, which provided important insight into the Young's modulus and their nonlinear deformation behavior [20-22]. However, most of these studies did not yet incorporate biochemical and molecular details into their investigations. To the best of the authors' knowledge, exceptions are the pioneering works by Sasaki and Odajima [23] and those by Fratzl *et al.* [24]. By applying x-ray diffraction methods these groups investigated the elongation mechanism of tendon collagen on the basis of the hierarchical structure of the tissue, and including the arrangement of collagen molecules in the tissue. They proposed models to describe how collagen molecules in fibrils are elongated and rearranged due to external force. In particular, three major deformation mechanisms were proposed; elongation of individual collagen molecules, an increase in the gap region between longitudinally adjoining molecules along the fibril axis, and relative slippage of laterally adjoining molecules along the fibril axis [24, 25].

To complement experimental approaches, molecular modeling provides a powerful approach to describe the molecular mechanics of collagen. Most earlier studies, however, were based on ultra-short collagen-like peptides obtained from x-ray crystallography [26-29]. The early molecular simulation studies used these short collagen molecules [30-39] that were typically limited to less than

10 nm length or more than a factor of 30 smaller than actual molecules found in collagen tissues. The resulting elastic modulus of these short collagen peptides was found to be on the order of 4.8±2 GPa, and much greater than the typical Young's moduli measured for macro-scale collagen tissues but in agreement with single molecule studies [16, 17, 40-42] (see Table 1 for an overview). The direct study of larger assemblies of collagen molecules into microfibrils and fibers with full atomistic simulation methods has remained elusive, due to the lack of an appropriate atomistic description and the size of the system. Previous reports of molecular modeling of collagen microfibrils based on a two-dimensional coarse-grained model, in which collagen molecules are described in a mesoscale bead-spring model [43, 44], predicted mechanical properties that did not match quantitatively with experimental findings. While the bead-spring model showed several features of the stress-strain behavior found in experiments, a disagreement of the magnitude of the predicted modulus was identified that could be reconciled. Furthermore, the earlier bead-spring model retained little information of the primary sequence, did not include a description of the three-dimensional arrangement of collagen molecules and lacked the ability to deal with explicit water solvent. These issues are, however, likely important for collagen mechanics and must be incorporated in a rigorous bottom-up tissue mechanics description.

In order to determine how collagen-based structures confer mechanical properties to tissues like skin, tendon and bone, and to identify how cells interact with the ECM, the understanding of the mechanics at different hierarchical levels and their interplay from a biochemical and molecular level upwards is essential. Earlier work has demonstrated that mechanical strain is distributed over distinct hierarchical levels (molecules, fibrils, fibres) [15, 45, 46] and that collagen tissue stretching involves concurrent deformation mechanisms. The significance of defining the material properties of collagenous tissues from the biochemistry level upwards is evident when considering the effect of mutations in collagen, which can result in incorrectly folded collagen protein that cause a variety of severe and sometimes deadly pathologies, such as Ehlers-Danlos syndrome or *osteogenesis imperfecta* (brittle bone disease) [47].

Results and discussion

Here we use an atomistic collagen microfibril model that includes full-length molecules with the actual amino acid sequence defined by the collagen gene and that thus completely captures the biochemical features of collagen molecules to describe the mechanical behavior at the microfibril level (see Materials and Methods for details). To the best of our knowledge, no such modeling of the mechanical properties at this scale has been previously attempted. The basis of our microfibril model is the recently reported structure of native *in situ* collagen in rat tail tendon [4, 6]. In pioneering experimental studies, by employing crystallographic techniques in x-ray fiber diffraction experiments, Orgel, Wess, *et al.* [6] obtained the packing arrangement of collagen molecules in a collagen microfibril, resulting in the three-dimensional geometry of collagen molecules including the N- and C-telopeptides. Based on the data from x-ray diffraction experiments, however, only the positions of the C_α backbone atoms in the collagen microfibril are available, and as a result the model reported in [6] did not yet include all atomistic details of the supermolecular assembly in collagen fibrils. In order to develop a full atomistic representation we use a computational approach to add all missing atoms including the side chains into the structure and identify the most stable configuration by using the all-atom CHARMM force field and a statistical structure identification approach (see Materials and Methods). Since the backbone structure is known, this homology modeling followed by extensive molecular equilibration provides a reliable estimate of the structure of the side chains.

The resulting atomistic model features full ≈300 nm long collagen molecules including the telopeptide domains attached at the ends of each molecule, and incorporates the complete three-

dimensional arrangement of collagen molecules arranged in a periodic unit cell (Fig. 2a). The unit cell dimension in the *Z*-axis corresponds to the length of the characteristic *D*-period observed for collagen fibrils. Thus, a full length collagen molecule spans five periodic cells in the *Z*-axis direction. Figure 2b shows the N-terminal portion of the original collagen molecule (in red) with four periodic images represented in gray, illustrating how the unit cell represents a model for the larger-scale molecular assembly into collagen microfibrils. Since the model uses periodic boundary conditions it resembles infinitely large collagen microfibrils in each dimension. The staggering of the molecules along the molecular axis leads to the well-known *D*-banding periodicity (Fig. 2c), while the molecules are arranged in a quasi-hexagonal pattern in the orthogonal direction where five molecules form the characteristic microfibril structure (Fig. 2d). Within each periodic cell, collagen molecules interdigitate with neighboring molecules to form a supertwisted right-handed microfibril. The characteristic banded structure of the equilibrated atomistic model of the collagen microfibril emerges naturally due to the three-dimensional structure of both single molecules and their assembly in the longitudinal and axial directions, and is found to be stable in our molecular model. Figure 2e illustrates a *D*-period with five molecular strands that form a collagen microfibril, showing the gap and overlap regions that arise due to the fact that one of the strands forming the microfibril is shorter than the *D*-period itself. The obtained *D*-banding reproduces experimental microscopy images of collagen fibrils well, owing to the fact that our molecular model is based on x-ray diffraction data and stable after molecular equilibration [48] (Fig. 3a-b).

We first report an account of the structural features of a fully equilibrated full-atomistic collagen microfibril in both hydrated (wet) and dehydrated (dry) conditions. In our study the dehydrated (dry) collagen microfibril model is used to assess the effect of hydration on the mechanical properties of collagen fibrils, which has been shown experimentally to be an important factor. The equilibration of the hydrated (wet) collagen microfibril (Fig. 3c) leads to a density of ≈1.19 g/cm^3, a value that is halfway between the density of water and the density of dehydrated (dry) collagen, which has been estimated at 1.34 g/cm^3 [49]. A Ramachandran analysis of the solvated system (Fig. 3e, left) shows that the collagen microfibril lies in a region of the diagram (Psi ≈150°, Phi ≈ -75°) that is characteristic of the polyproline II chain and thus of collagen-like peptides, in good agreement with earlier experimental structural studies [50]. The density of the dehydrated (dry) collagen microfibril (Fig. 3d) reaches a larger density, with a value of ≈1.29 g/cm^3. A Ramachandran analysis of the dehydrated (dry) collagen microfibril (Fig. 3e, right) shows that it also lies in a region of the diagram that is characteristic of collagen-like peptides (Psi ≈150°, Phi ≈ -75°); however, a broader range of dihedral angles is found indicating some level of molecular unfolding. This suggests that the loss of water results in a loss of structure at the molecular scale. This may indicate that water is indeed needed to keep the characteristic configuration of collagen molecules in microfibrils, in agreement with earlier work on collagen-like peptides at the single molecule scale where a greater level of disorder and loss of H-bonding was found in vacuum studies [30]. Nevertheless, more in-depth studies of this issue for example with Replica Exchange Molecular Dynamics, along with experimental validation, will be necessary, combined with further experimental work [51]. Overall our model shows good quantitative agreement with available experimental structural data of collagen microfibrils. This confirms that the molecular representation based on x-ray data is a reasonable starting point for the analysis of its mechanical properties.

We now test the mechanical properties of a hydrated (wet) and dehydrated (dry) collagen microfibril by applying constant stress boundary conditions along the fibril axis and monitoring the resulting strain at equilibrium. We test stresses in the range from 0 to 200 MPa, leading to the stress-strain behaviors shown in Fig. 4a. We find that hydrated (wet) collagen microfibrils feature two distinct deformation regimes. In the small-strain regime (<10%) the predicted Young's modulus is ≈300 MPa, while in the large-strain regime (>10%) the microfibril shows a severely increased tangent

stiffness, with a Young's modulus of ≈1.2 GPa. The results of nanomechanical testing of hydrated (wet) collagen microfibril are in good agreement with available experimental results obtained for the small strain regime based on several techniques such as x-ray diffraction [23], Atomic Force Microscopy (AFM) [8, 21] and the use of micro-electro-mechanical systems (MEMS) [20, 22, 52]. Figure 4b-d and Table 1 present a systematic comparison with a range of experimental data. It is noted that for the larger-strain regime there exists less experimental information and available results tend to be more scattered. For example in recent work, Shen *et al.* [22] showed a relatively large variability of collagen fibril behaviors at large deformation, which suggested either strain-hardening or strain-softening depending on the fibril investigated. A direct comparison of the mechanical properties of single collagen molecules versus that of collagen microfibrils suggests that the mechanical properties are strongly scale dependent. We find a severe change of the modulus when comparing a single collagen molecule to a collagen fibril, as shown in Fig. 4b and Table 1. A direct numerical comparison suggests a factor of 10-20 difference in the Young's modulus from several GPa for a single molecule to a few hundred MPa for collagen microfibrils, presenting a striking change of mechanical properties at different hierarchical levels. This finding agrees well with experimental data as is confirmed in Table 1. We also study the mechanical properties of dehydrated (dry) collagen fibrils to test the effect of water solvent on the collagen mechanical properties at the fibril level, which allows us to explore an effect that had earlier been investigated in experimental AFM studies [21]. Our simulations suggest that dehydrated (dry) collagen microfibrils show an almost perfect linear elastic behavior, albeit with a much greater Young's modulus of ≈1.8 GPa (approaching ≈2.25 GPa for larger strains), or a striking factor of 6.75 larger than the modulus of hydrated (wet) collagen microfibril. A similar ratio of the moduli in dehydrated (dry) vs. hydrated (wet) states has been observed in experiment [21], where a modulus ratio of 9 between the dehydrated (dry) vs. hydrated (wet) states of the has been identified (see Table 1). These findings point to the great importance of water for the mechanics of collagen microfibrils, and raise an issue that should be further investigated using experimental and theoretical methods.

Our atomistic model enables us to observe atomistic and molecular deformation mechanisms not directly accessible to experimental techniques, and thereby to explain the molecular origin of mechanical properties at different hierarchical levels and levels of strain. We investigate the molecular mechanisms during fibril stretching, aimed at elucidating the mechanisms behind the two regimes observed in the stress-strain curve (Fig. 4a), and the mechanical properties in hydrated (wet) and dehydrated (dry) conditions. For the hydrated (wet) case, the small deformation the collagen molecules' end-to-end distance increases linearly until the microfibril strain reaches 10% (corresponding to ≈50 MPa stress), the strain at which the microfibril stiffness increases drastically. Beyond this point the molecular end-to-end still continues to increase but the slope of curve is significantly lower (Fig. 5a). This can be explained due to the fact that below 10% strain the collagen molecule is straightened within the microfibril and thereby loses its kinked arrangement, while beyond the 10% strain mark the molecule itself is being stretched, resulting in a larger mechanical resistance to deformation. The monitoring of the dihedral energy of the systems confirms this hypothesis and shows that for strains larger than 10% the dihedral energy increases, which directly confirms that the molecule is being deformed (Fig. 5b).

The increase in the gap to overlap ratio in the small-strain regime (Fig. 5c) suggests that the initial straightening is concentrated in the gap regions where the molecular packing density is lower and molecules are less organized, with more kinks. Our model thereby directly confirms a suggestion made by Fratzl *et al*. [24, 53, 54] (Fig. 5d). However, it is noted that the straightening of collagen molecules is not limited to the gap regions as collagen molecules in the microfibril feature a kinked geometry throughout the entire structure, which is successively lost during deformation. Once the entire capacity to molecular straightening is exhausted and all collagen molecules assume a straight

configuration oriented in the direction of the pulling axis, collagen molecules themselves undergo stretching, leading to a significantly increased tangent microfibril stiffness, at strains in excess of 10%. The combination of these two mechanisms, molecular straightening and molecular stretching, effectively lead to an increase of the *D*-period, in agreement with experimental results [55]. Other mechanisms, such as molecular sliding may take place at larger strains in excess of 30% as described in earlier studies of the deformation mechanisms [23, 24]. Conversely, for the dehydrated (dry) collagen microfibril the deformation mechanisms in the investigated stress range involves primarily the straightening of densely packed collagen molecules without stretching of the molecule itself. This is shown by the increasing molecular end-to-end distance (Fig. 5a), which increases linearly, and by the dihedral energy (Fig. 5b), which does not increases with the strain. The analysis of the gap-to-overlap ratio (Fig. 5c) further shows that deformation is initially equally distributed in both the gap and overlap regions (where the ratio remains constant), and that for larger stresses deformation increasingly affects the gap region (shown by the increase in the gap-to-overlap ratio).

Conclusion

In summary, here we used a three-dimensional full-atomistic collagen microfibril model to describe the mechanics of collagen microfibrils from a bottom-up perspective. Our model includes full-length collagen molecules with a complete representation of the entire amino acid sequence, full atomistic details, as well as explicit water solvent. The model, based on experimental data of an *in situ* collagen molecule [6] and equilibrated through full-atomistic molecular dynamics simulations, captures all major structural features of collagen microfibrils such as the quasi-hexagonal molecular packing, the *D*-banding periodicity (Fig. 2), the distribution of dihedral angles (Fig. 3) as well as the broad range of mechanical behaviour (Fig. 4). All structural and mechanical data agrees well with experimental results from various sources, as summarized in Table 1. Our results suggest that deformation of collagen microfibrils is mediated through mechanisms that operate at different hierarchical levels, involving straightening of disordered and helically twisted molecules at small strains, first in the gap regions and then in the entire fibril, followed by axial stretching of molecules, and eventual molecular uncoiling (Figs. 4 and 5).

The observation of these deformation mechanisms directly confirm hypotheses put forth in earlier experimental works from a biochemical molecular perspective. Moreover, these mechanisms explain the striking difference of the Young's modulus of collagen microfibril compared with that of single molecules, which is typically found in the range of 4.8±2 GPa or ≈10-20 times greater than that of a collagen microfibril. This resolves a long-standing issue in collagen mechanics modelling that has thus far prevented the consolidation of experimental findings with earlier computational results [36, 43, 44]. Our findings also substantiate the notion that the properties of collagen tissues are highly dependent on the hierarchical level and deformation state (*i.e.*, strain) considered. This suggests that many conventional continuum models of collagen tissues may not be adequate to describe the complex scale-dependent and nonlinear mechanical properties. Future work could be integrated with recent theoretical studies on the effect of hierarchical mechanisms in bone-like materials, and provide a computational validation for predictions made about the role of different hierarchy levels in tissues such as bone [56].

A key impact of our results is that the experimentally validated molecular model of collagen microfibril mechanics provides a basis to investigate collagenous tissues at the fibril scale and larger, which could change the paradigm of collagen research. Indeed, with our model it is possible to assess, from a bottom-up perspective, how changes at the biochemical and atomistic level (such as amino acid mutations, cross-link patterns or density) affect the structural and mechanical properties at the microscale and macroscale. This, together with the study of the interaction of collagen fibrils with other relevant biomolecules such as proteoglycans, will provide critical details for the understanding

of structure-property relationships in the broader class of collagen materials, and at larger hierarchical scales. Indeed, important experimental crystallography [57, 58] efforts are ongoing to elucidate the structure of collagen fibril-proteoglycan complexes. Other challenges remain with respect to the greater level of disorder that is expected to be found in collagen fibrils, as outlined in [59]. Our model represents a collagen "microfibril", whereas larger-scale collagen fibrils may feature additional interfaces and disorder between them that could affect the mechanical properties. The construction of such a full fibril mechanics model could be addressed in future work, perhaps via a combination of experimental and computational techniques. However, computational challenges associated with such modeling are daunting as the construction of such a model would involve billions of atoms, a size that is currently out of reach for protein simulations.

Our results could be important in the context of mechanical properties of collagen tissues for cell culture and scaffolding material development. It has been shown that a cell's microenvironment is important in stem cell lineage specification, where soft matrices that resemble brain tissue-like moduli are neurogenic, stiffer matrices that mimic muscle are myogenic, and rigid matrices that mimic bone prove to be osteogenic [60]. Cells act at the micron scale, the scale of collagen microfibrils, fibrils and fibers, and their behavior is probably affected by the complex hierarchical structure of their surrounding environment. However, current biomaterials used for scaffolding do not present a hierarchical structure as that found in natural materials and this could affect the cell behavior and differentiation. A rigorous understanding of collagen's scale and strain dependent stiffness may help in designing biomaterials with appropriate mechanical characteristics and thus addresses an immediate need for optimized matrix elasticity to foster differentiation and regeneration for regenerative medicine applications based on stem cell therapies such as cardiomyoplasty, muscular dystrophy, and neuroplasty [61-63]. To highlight the variation of Young's modulus and bending rigidity for a variety of biological and synthetic fibers we present a comparative analysis as shown in Fig. 6. This analysis demonstrates that collagen fibrils provide a significant bending rigidity at relatively high Young's modulus.

The model of collagen microfibril mechanics reported here is based on x-ray diffraction results and is limited by the range of molecular conformation changes that can be observed at a molecular dynamics time-scale. For example, even though the mechanical analysis of the modulus of dehydrated (dry) collagen microfibril agrees well with experimental findings (with similar Young's modulus values as shown in Table 1), the Ramachandran analysis suggest a heightened level of disorder in the system, perhaps indicative of molecular unfolding. Drastic changes in the molecular architecture associated with such mechanisms could be explored via the use of advanced computational methods, such as Replica Exchange Molecular Dynamics and should be combined with experimental efforts such as outlined in [59]. The computational challenges associated with these methods are, however, enormous. A limitation of the collagen microfibril mechanics model reported here is that it is based on the periodic repetition of a crystallographic unit cell, due to the significant computational cost associated with simulating this large molecular structure. The periodic model also implies that no cross-links between molecules are considered, and that sliding between molecules is not taken into account. Despite these limitations, as confirmed in Table 1 the model properly captures the mechanical behavior seen in experiments, likely because the above listed constraints do not affect the behavior at relatively small deformation, below 20-30% strain. Indeed, the effects of cross-links between molecules and intermolecular sliding have been shown to dominate the properties primarily at larger deformation [64]. A microfibril model that explicitly considers multiple molecules poses no fundamental challenge; however, it would be rather challenging from a computational point of view. As appropriate computational resources become available, however, development of such models will be straightforward by extending the work reported here. The use of a coarse-graining approach as applied in previous works [65] may be helpful to reduce the

computational cost and may allow us to model a collagen microfibril that explicitly considers multiple molecules, cross-links between them, and even the interactions with proteoglycans [57, 58] and cellular components such as integrins [66]. However, such a coarse-grained model will need to be carefully validated against experimental and full-atomistic simulation results.

Materials and methods

Existing collagen microfibril and fibril models represent the supramolecular arrangement in collagenous tissues in a simplified way, using a two-dimensional lattice of mesoscopic beads [43, 44] or extremely short collagen peptides [30-39]. These models do not account for the biochemical details and are much smaller than the typical length-scales of collagen molecules found in collagen microfibrils and fibrils. The reason for these approximations is that up until now crystallographic details (and in particular a full-atomistic geometry) of the collagen molecule have been obtained only for short collagen-like peptides with lengths below ≈10 nm [26-28]. The method applied here overcomes these limitations and enables us to develop a full atomistic model of the mechanics of collagen microfibrils.

***Homology modelling*:** The structural model of the collagen microfibril is generated starting from the *in situ* structure of full length collagen type I molecule [6] (Protein Data Bank identification code 3HR2). This structure, obtained by employing conventional crystallographic techniques in x-ray fiber diffraction experiments, resolved for the first time the specific three-dimensional arrangement of collagen molecules in naturally occurring fibrils, including the N- and C-telopeptides. Since the structure reported in [6] includes only backbone alpha carbons and the primary sequence of *rattus norvegicus*, we used homology modelling to obtain a full-atom structure with the human collagen sequence. The sequence of the human type I collagen is obtained from PubMed (entry number NP_000079 for α1(I) chain and NP_000080 for the α2(I) chain). The 3HR2 structure and the human collagen type I sequence are aligned, and ten homology models are built and scored by the Discrete Optimized Protein Energy (DOPE) using the Modeller program (version 9.6) [67]. The structure with the lowest DOPE value is chosen for building the collagen microfibril model.

Fibril model generation: The collagen supramolecular model (microfibril) is generated using the information on the naturally occurring crystallographic unit cell reported in [6], and in the associated 3HR2 structure (a≈ 40.0 Å, b≈27.0 Å, c≈678 Å, α≈89.2°, β≈94.6°, γ≈105.6°). The molecular packing topology obtained by the periodic repetition of the unit cell lead to quasi-hexagonally packed collagen molecules which interdigitates with neighboring molecules to form a supertwisted right-handed microfibril and to the well known *D*-banding periodicity seen in AFM images of collagen fibrils. The fibril model is solvated using the solvate plug in of GROMACS by adding SPC water molecules. Since the molecule at physiological pH includes a net charge (positive net charge +34), counterions (Cl^-) are added in order to keep the system neutral. The final solvated all-atom system contains ≈57,000 atoms, including ≈32,000 water atoms. The first step of energy minimization is performed by a steepest descent algorithm using the GROMACS 4.0 code [68, 69] and the GROMOS 43a1 force field, which includes parameters for hydroxyproline amino acid (HYP) found in collagen. This force field has been widely validated for a variety of biochemical models of proteins including collagen [34-38].

All-atom equilibration: Full atomistic simulations are carried out using the GROMACS 4.0 code [68, 69]. Rigid bonds are used to constrain covalent bond lengths, thus allowing an integration time step of 2 fs. Nonbonding interactions are computed using a cut-off for neighbor list at 1.35 nm, with a switching function between 1.0 and 1.2 nm for Van Der Waals interactions, while the Particle-Mesh Ewald sums (PME) method is applied to describe electrostatic interactions. The fibril model is equilibrated through 8.5 ns NPT molecular dynamics simulations at a temperature of 310 K (37°C),

and with 1 bar pressure (we use velocity-rescaling thermostat with 1 ps coupling constant and a Berendsen barostat with 1 ps time constant). This ensures structural convergence through a Root Mean Square Deviation (RMSD) analysis, where convergence is confirmed when the slope of the RMSD with respect to time approaches zero.

In silico mechanical testing: In order to assess the mechanical properties of the hydrated and dehydrated atomistic microfibril models we perform molecular dynamics simulations with increasing constant mechanical stress in tension along the fibril axis while maintaining the pressure on the other axes constant at 1 bar (using a Berendsen barostat and 1 ps coupling constant). Although the *c* axis (the long axis) of the periodic unit cell is not perfectly aligned parallel with the microfibril direction (by an angle mismatch of about 1 degree), the effect on the mechanical properties (the focus of this study) is very small since the mechanical load is reasonably well aligned with the overall direction of the microfibril. The mechanical loading implemented here reflects the setup that is also used for mechanical testing in experimental studies. The applied stresses are in the range from 0 to 200 MPa, applied during 20 ns molecular dynamics simulation for each load applied. We find that equilibrium of the molecular structure is reliably reached within 10-15 ns of molecular dynamics simulation, depending on the extent of the deformation. Thus we use the last 5 ns of molecular dynamics simulation for the mechanical analysis after the system has fully converged. To ensure that equilibrium is obtained, we monitor pressure equilibrium, protein RMSD and confirm that the size of the simulation cell reaches a steady value. The strain $\varepsilon(\sigma)$ is calculated as follows:

$$\varepsilon(\sigma) = \frac{L(\sigma) - L_0}{L_0}, \qquad (1)$$

where $L(\sigma)$ is the equilibrium cell length along the microfibril axis when a strain σ is applied, while L_0 is the equilibrium cell length along the fibril axis for σ=0. From the fibril strains $\varepsilon(\sigma)$ resulting from each applied stress σ we obtain the stress-strain behaviour as plotted in Fig. 4.

Computational method and cost: Due to the size of the model, all-atom simulations of the collagen microfibril with explicit solvent are computationally very intense. The solvated (wet) full-atomistic model contains ≈57,000 atoms (≈25,000 in the dehydrated [dry] model), requiring about 6 hours per nanosecond on 32 CPUs on a parallel machine. Since phenomena involving the molecular (and supramolecular) scale are often in the range of several nanoseconds or even microseconds, molecular dynamics simulations of the full-atomistic collagen fibril are at the edge of current computational capabilities.

Structures in PDB format: Geometries of all relevant structures in PDB format are available from the corresponding author upon request.

Acknowledgements: The authors thank Professor Joseph Orgel (Pritzker Institute of Biomedical Science and Engineering at the Illinois Institute of Technology and Argonne National Lab) and Professor Sandra Shefelbine (Department of Bioengineering at Imperial College, London) for their helpful suggestions during the preparation of this manuscript. This research was supported by NSF (grant # CMMI-0642545), ONR (grant # N000141010562) and by the MIT-Italy program "Progetto Rocca". We acknowledge CINECA (Award IscrB TMC-OI, 2010) and CILEA (Award TMC-OI, 2010) for the provision of high performance computing resources and support.

Competing interests statement: The authors declare no competing interests of any sort.

References

1. Kadler, K.E., C. Baldock, J. Bella, and R.P. Boot-Handford, *Collagens at a glance.* Journal of Cell Science, 2007. **120**(12): p. 1955-1958.
2. Fraser, R.D., T.P. MacRae, A. Miller, and E. Suzuki, *Molecular conformation and packing in collagen fibrils.* J Mol Biol, 1983. **167**(2): p. 497-521.
3. Hulmes, D.J.S., T.J. Wess, D.J. Prockop, and P. Fratzl, *Radial Packing, Order, And Disorder In Collagen Fibrils.* Biophysical Journal, 1995. **68**(5): p. 1661-1670.
4. Orgel, J.P.R.O., A. Miller, T.C. Irving, R.F. Fischetti, A.P. Hammersley, and T.J. Wess, *The In Situ Supramolecular Structure of Type I Collagen.* Structure, 2001. **9**: p. 1061-1069.
5. Currey, J.D., *Bones: Structure and Mechanics*. 2002, Princeton, NJ: Princeton University Press.
6. Orgel, J.P.R.O., T.C. Irving, A. Miller, and T.J. Wess, *Microfibrillar structure of type I collagen in situ.* P. Natl. Acad. Sci. USA, 2006. **103**(24): p. 9001-9005.
7. Fratzl, P., *Collagen: Structure and Mechanics*. 2008: Springer (New York).
8. Aladin, D.M., K.M. Cheung, A.H. Ngan, D. Chan, V.Y. Leung, C.T. Lim, K.D. Luk, and W.W. Lu, *Nanostructure of collagen fibrils in human nucleus pulposus and its correlation with macroscale tissue mechanics.* J Orthop Res. **28**(4): p. 497-502.
9. Petruska, J.A. and A.J. Hodge, *Subunit Model for Tropocollagen Macromolecule.* Proceedings of the National Academy of Sciences of the United States of America, 1964. **51**(5): p. 871-&.
10. Moeller, H.D., U. Bosch, and B. Decker, *Collagen Fibril Diameter Distribution in Patellar Tendon Autografts after Posterior Cruciate Ligament Reconstruction in Sheep - Changes over Time.* Journal of Anatomy, 1995. **187**: p. 161-167.
11. Weiner, S. and H.D. Wagner, *The material bone: Structure mechanical function relations.* Annual Review Of Materials Science, 1998. **28**: p. 271-298.
12. Fratzl, P. and R. Weinkamer, *Nature's hierarchical materials.* Progress in Materials Science, 2007. **52**: p. 1263-1334.
13. Camacho, N.P., L. Hou, T.R. Toledano, W.A. Ilg, C.F. Brayton, C.L. Raggio, L. Root, and A.L. Boskey, *The material basis for reduced mechanical properties in oim bones.* Journal of Bone and Mineral Research 1999. **14**(2): p. 264-272.
14. Fratzl, P., H.S. Gupta, E.P. Paschalis, and P. Roschger, *Structure and mechanical quality of the collagen-mineral nano-composite in bone.* Journal Of Materials Chemistry, 2004. **14**(14): p. 2115-2123.
15. Gupta, H.S., J. Seto, W. Wagermaier, P. Zaslansky, P. Boesecke, and P. Fratzl, *Cooperative deformation of mineral and collagen in bone at the nanoscale.* P. Natl. Acad. Sci. USA, 2006. **103**: p. 17741-17746.
16. Cusack, S. and A. Miller, *Determination Of The Elastic-Constants Of Collagen By Brillouin Light-Scattering.* Journal Of Molecular Biology, 1979. **135**(1): p. 39-51.
17. Sun, Y., Z. Luo, A. Fertala, and K. An, *Direct quantification of the flexibility of type I collagen monomer.* Biochemical And Biophysical Research Communications, 2002. **295**(2): p. 382-386.
18. Bozec, L. and M. Horton, *Topography and mechanical properties of single molecules of type I collagen using atomic force microscopy.* Biophysical Journal, 2005. **88**(6): p. 4223-4231.
19. Buehler, M.J. and S.Y. Wong, *Entropic elasticity controls nanomechanics of single tropocollagen molecules.* Biophysical Journal, 2007. **93**(1): p. 37-43.
20. Eppell, S.J., B.N. Smith, H. Kahn, and R. Ballarini, *Nano measurements with micro-devices: mechanical properties of hydrated collagen fibrils.* Journal Of The Royal Society Interface, 2006. **3**(6): p. 117-121.

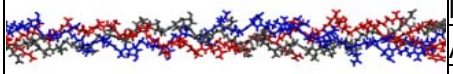
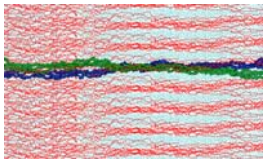
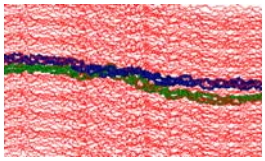
21. van der Rijt, J.A.J., K.O. van der Werf, M.L. Bennink, P.J. Dijkstra, and J. Feijen, *Micromechanical testing of individual collagen fibrils.* Macromolecular Bioscience, 2006. **6**(9): p. 697-702.
22. Shen, Z.L., M.R. Dodge, H. Kahn, R. Ballarini, and S.J. Eppell, *Stress-strain experiments on individual collagen fibrils.* Biophysical Journal, 2008. **95**(8): p. 3956-3963.
23. Sasaki, N. and S. Odajima, *Elongation mechanism of collagen fibrils and force-strain relations of tendon at each level of structural hierarchy.* Journal Of Biomechanics, 1996. **29**(9): p. 1131-1136.
24. Fratzl, P., K. Misof, I. Zizak, G. Rapp, H. Amenitsch, and S. Bernstorff, *Fibrillar structure and mechanical properties of collagen.* Journal of Structural Biology, 1998. **122**(1-2): p. 119-122.
25. Fratzl, P., I. Burgert, and H.S. Gupta, *On the role of interface polymers for the mechanics of natural polymeric composites.* Physical Chemistry Chemical Physics, 2004. **2004**: p. 5575-5579.
26. Beck, K., V.C. Chan, N. Shenoy, A. Kirkpatrick, J.A.M. Ramshaw, and B. Brodsky, *Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine.* P. Natl. Acad. Sci. USA, 2000. **97**(8): p. 4273-4278.
27. Bella, J., M. Eaton, B. Brodsky, and H.M. Berman, *Crystal-structure and molecular-structure of a collagen-like peptide at 1.9-angstrom resolution.* Science, 1994. **266**(5182): p. 75-81.
28. Xu, Y., A. Persikov, J. Jordan, and B. Brodsky, *Thermodynamic analysis of collagen-like peptides.* Biophys. J., 2000. **78**(1): p. 425A-425A.
29. Redaelli, A., S. Vesentini, M. Soncini, P. Vena, S. Mantero, and F.M. Montevecchi, *Possible role of decorin glycosaminoglycans in fibril to fibril force transfer in relative mature tendons - a computational study from molecular to microstructural level.* Journal Of Biomechanics, 2003. **36**(10): p. 1555-1569.
30. Mogilner, I.G., G. Ruderman, and J.R. Grigera, *Collagen stability, hydration and native state.* Journal of Molecular Graphics & Modelling, 2002. **21**(3): p. 209-213.
31. Lorenzo, A.C. and E.R. Caffarena, *Elastic properties, Young's modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics.* Journal Of Biomechanics, 2005. **38**(7): p. 1527-1533.
32. Buehler, M.J., *Atomistic and continuum modeling of mechanical properties of collagen: Elasticity, fracture and self-assembly.* J. Mater. Res., 2006. **21**(8): p. 1947-1961.
33. Bhowmik, R., K.S. Katti, and D.R. Katti, *Mechanics of molecular collagen is influenced by hydroxyapatite in natural bone.* Journal of Materials Science, 2007. **42**(21): p. 8795-8803.
34. Gautieri, A., S. Vesentini, F.M. Montevecchi, and A. Redaelli, *Mechanical properties of physiological and pathological models of collagen peptides investigated via steered molecular dynamics simulations.* Journal Of Biomechanics, 2008. **41**(14): p. 3073-3077.
35. Gautieri, A., M.J. Buehler, and A. Redaelli, *Deformation rate controls elasticity and unfolding pathway of single tropocollagen molecules.* Journal of the Mechanical Behavior of Biomedical Materials, 2009. **2**(2): p. 130-137.
36. Gautieri, A., S. Uzel, S. Vesentini, A. Redaelli, and M.J. Buehler, *Molecular and Mesoscale Mechanisms of Osteogenesis Imperfecta Disease in Collagen Fibrils.* Biophysical Journal, 2009. **97**(3): p. 857-865.
37. Gautieri, A., S. Vesentini, A. Redaelli, and M.J. Buehler, *Single molecule effects of osteogenesis imperfecta mutations in tropocollagen protein domains.* Protein Science, 2009. **18**(1): p. 161-168.
38. Srinivasan, M., S.G.M. Uzel, A. Gautieri, S. Keten, and M.J. Buehler, *Alport Syndrome mutations in type IV tropocollagen alter molecular structure and nanomechanical properties.* Journal of Structural Biology, 2009. **168**(3): p. 503-510.

39. Bhowmik, R., K.S. Katti, and D.R. Katti, *Mechanisms of Load-Deformation Behavior of Molecular Collagen in Hydroxyapatite-Tropocollagen Molecular System: Steered Molecular Dynamics Study.* Journal of Engineering Mechanics-Asce, 2009. **135**(5): p. 413-421.
40. Hofmann, H., T. Voss, K. Kuhn, and J. Engel, *Localization Of Flexible Sites In Thread-Like Molecules From Electron-Micrographs - Comparison Of Interstitial, Basement-Membrane And Intima Collagens.* Journal Of Molecular Biology, 1984. **172**(3): p. 325-343.
41. Sasaki, N. and S. Odajima, *Stress-strain curve and Young's modulus of a collagen molecule as determined by the X-ray diffraction technique.* Journal of Biomechanics, 1996. **29**(5): p. 655-658.
42. Harley, R., D. James, A. Miller, and J.W. White, *Phonons And Elastic-Moduli Of Collagen And Muscle.* Nature, 1977. **267**(5608): p. 285-287.
43. Buehler, M.J., *Nature designs tough collagen: Explaining the nanostructure of collagen fibrils.* P. Natl. Acad. Sci. USA, 2006. **103**(33): p. 12285–12290.
44. Buehler, M.J., *Nanomechanics of collagen fibrils under varying cross-link densities: Atomistic and continuum studies.* Journal of the Mechanical Behavior of Biomedical Materials, 2008. **1**(1): p. 59-67.
45. Gupta, H.S., W. Wagermaier, G.A. Zickler, D.R.B. Aroush, S.S. Funari, P. Roschger, H.D. Wagner, and P. Fratzl, *Nanoscale deformation mechanisms in bone.* Nano Letters, 2005. **5**(10): p. 2108-2111.
46. Gupta, H.S., P. Messmer, P. Roschger, S. Bernstorff, K. Klaushofer, and P. Fratzl, *Synchrotron diffraction study of deformation mechanisms in mineralized tendon.* Physical Review Letters, 2004. **93**(15).
47. Rauch, F. and F.H. Glorieux, *Osteogenesis Imperfecta.* The Lancet, 2004. **363**: p. 1377-85.
48. Holmes, D.F., C.J. Gilpin, C. Baldock, U. Ziese, A.J. Koster, and K.E. Kadler, *Corneal collagen fibril structure in three dimensions: Structural insights into fibril assembly, mechanical properties, and tissue organization.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(13): p. 7307-7312.
49. Chapman, E.W. and F. Rodriguez, *Acrylic resin reinforcement of reconstituted collagen films.* Polymer Engineering and Science, 1977. **17**(5): p. 282-286.
50. Brodsky, B. and J.A.M. Ramshaw, *The collagen triple-helix structure.* Matrix Biology, 1997. **15**(8-9): p. 545-554.
51. Mizuno, K. and H.P. Bachinger, *The effect of deuterium oxide on the stability of the collagen model peptides H-(Pro-Pro-Gly)(10)-OH, H-(Gly-Pro-4(R)Hyp)(9)-OH, and Type I collagen.* Biopolymers. **93**(1): p. 93-101.
52. Espinosa, H.D., Y. Zhu, and N. Moldovan, *Design and operation of a MEMS-based material testing system for nanomechanical characterization.* Journal Of Microelectromechanical Systems, 2007. **16**(5): p. 1219-1231.
53. Fratzl, P., ed. *Collagen: Structure and Mechanics*. 2008, Springer.
54. Fratzl, P. and R. Weinkamer, *Nature's hierarchical materials.* Progress in Material Science, 2007. **52**: p. 1263-1334.
55. Kukreti, U. and S.M. Belkoff, *Collagen fibril D-period may change as a function of strain and location in ligament.* Journal of Biomechanics, 2000. **33**(12): p. 1569-1574.
56. Zhang, Z., Y.W. Zhang, and H. Gao, *On optimal hierarchy of load-bearing biological materials.* Proc. Roy. Soc. B, 2010.
57. Orgel, J.P., O. Antipova, I. Sagi, A. Bitler, D. Qiu, R. Wang, Y. Xu, and J. San Antonio, *Collagen fibril surface displays a constellation of sites capable of promoting fibril assembly, stability, and homeostasis.* Connective Tissue Research, in press.
58. Orgel, J.P., J. San Antonio, and O. Antipova, *Molecular and structural mapping of collagen fibril interaction.* Connective Tissue Research, in press.

59. Perumal, S., O. Antipova, and J.P. Orgel, *Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis.* Proc Natl Acad Sci U S A, 2008. **105**(8): p. 2824-9.
60. Engler, A.J., S. Sen, H.L. Sweeney, and D.E. Discher, *Matrix elasticity directs stem cell lineage specification.* Cell, 2006. **126**(4): p. 677-689.
61. Yow, S.Z., C.H. Quek, E.K. Yim, C.T. Lim, and K.W. Leong, *Collagen-based fibrous scaffold for spatial organization of encapsulated and seeded human mesenchymal stem cells.* Biomaterials, 2009. **30**(6): p. 1133-42.
62. LeDuc, P. and R. Schwartz, *Computational models of molecular self-organization in cellular environments.* Cell Biochem Biophys, 2007. **48**(1): p. 16-31.
63. LeDuc, P.R. and D.N. Robinson, *Using Lessons from Cellular and Molecular Structures for Future Materials.* Advanced Materials, 2007. **19**: p. 3761-3770.
64. Uzel, S. and M.J. Buehler, *Molecular structure, mechanical behavior and failure mechanism of the C-terminal cross-link domain in type I collagen.* Journal of the Mechanical Behavior of Biomedical Materials, 2010. in press, doi:10.1016/j.jmbbm.2010.07.003.
65. Gautieri, A., A. Russo, S. Vesentini, A. Redaelli, and M.J. Buehler, *Coarse-Grained Model of Collagen Molecules Using an Extended MARTINI Force Field.* Journal of Chemical Theory and Computation, 2010. **6**(4): p. 1210-1218.
66. Kumar, S. and P.R. Leduc, *Dissecting the Molecular Basis of the Mechanics of Living Cells.* Experimental Mechanics, 2009. **49**(1): p. 11-23.
67. Fiser, A. and A. Sali, *MODELLER: Generation and refinement of homology-based protein structure models.* Macromolecular Crystallography, Pt D, 2003. **374**: p. 461-+.
68. Berendsen, H.J.C., D. Vanderspoel, and R. Vandrunen, *Gromacs - a message-passing parallel molecular-dynamics implementation.* Computer Physics Communications, 1995. **91**(1-3): p. 43-56.
69. Vandrunen, R., D. Vanderspoel, and H.J.C. Berendsen, *Gromacs - a software package and a parallel computer for molecular-dynamics.* Abstracts of Papers of the American Chemical Society, 1995. **209**: p. 49.
70. Vesentini, S., C.F.C. Fitie, F.M. Montevecchi, and A. Redaelli, *Molecular assessment of the elastic properties of collagen-like homotrimer sequences.* Biomechanics And Modeling In Mechanobiology, 2005. **3**(4): p. 224-234.
71. Smith, J.F., T.P.J. Knowles, C.M. Dobson, C.E. MacPhee, and M.E. Welland, *Characterization of the nanoscale properties of individual amyloid fibrils.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(43): p. 15806-15811.
72. Nishino, T., R. Matsui, and K. Nakamae, *Elastic modulus of the crystalline regions of chitin and chitosan.* Journal of Polymer Science Part B-Polymer Physics, 1999. **37**(11): p. 1191-1196.
73. Gosline, J.M., P.A. Guerette, C.S. Ortlepp, and K.N. Savage, *The mechanical design of spider silks: From fibroin sequence to mechanical function.* Journal of Experimental Biology, 1999. **202**(23): p. 3295-3303.
74. Keten, S., Z.P. Xu, B. Ihle, and M.J. Buehler, *Nanoconfinement controls stiffness, strength and mechanical toughness of beta-sheet crystals in silk.* Nature Materials, 2010. **9**(4): p. 359-367.
75. Deriu, M.A., S. Enemark, M. Soncini, F.M. Montevecchi, and A. Redaelli, *Tubulin: from atomistic structure to supramolecular mechanical properties.* Journal Of Materials Science, 2007. **42**(21): p. 8864-8872.

Tables and Table Legends

Table 1 | Comparison of Young's modulus of collagen molecule (solvated) and collagen microfibril (including hydrated [wet] and dehydrated [dry] states) as predicted from experimental and theoretical analyses.

Hierarchical level or state	Type of analysis	Young's Modulus (GPa)
Collagen molecule (averaged value 4.8±2.0 GPa) 	X-Ray diffraction [41]	≈3
	Brillouin light scattering [16]	≈5.1
	Brillouin light scattering [42]	≈9
	Estimate based on persistence length [17]	0.35-12
	Estimate based on persistence length [40]	≈3
	Atomistic modeling [31]	≈4.8
	Reactive atomistic modeling [32]	≈7
	Atomistic modeling [70]	≈2.4
	Atomistic modeling [35]	≈4
	Coarse-grain modeling [65]	≈4
Collagen microfibril, wet (averaged value 0.6±0.2 GPa) 	X-ray diffraction [23]	0.43 (small strain)
	AFM testing [21]	0.2-0.8 (small strain)
	MEMS stretching [20, 22]	0.86 ± 0.45 (small strain) 0.72 ± 0.57 (large strain)
	Bead-spring based mesoscale model [43, 44]	4.36 (smaller strain) ≈38 (significantly larger strain)
	Atomistic modeling *(present study)*	≈ 0.3 (small strain) ≈ 1.2 (large strain)
Collagen microfibril, dry (averaged value 3.26 GPa) 	AFM testing [21]	2-7
	Atomistic modeling *(present study)*	≈ 1.8 ≈ 2.25 (larger strain)

Figures and Figure Legends

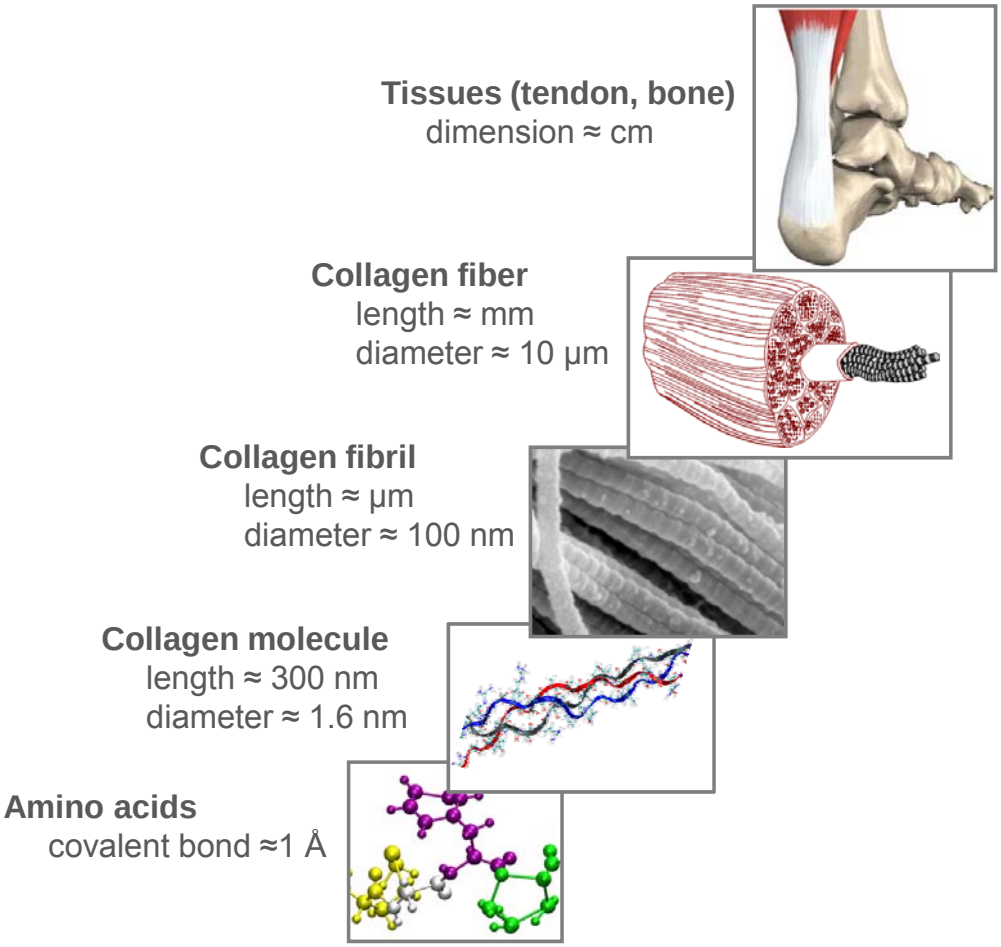


Figure 1 | Hierarchical structure of collagen protein materials. Each collagen molecule is made of three peptide chains that form the ≈300 nm long triple helical collagen molecule. Collections of collagen molecules aggregate both in lateral and longitudinal directions to form fibrils. Fibrils in cornea are normally thin (≈30 nm) and uniform in diameter, while tissues such as tendon contain a wide-ranging distribution of diameters (100-500 nm). Fibrils include tiny hydroxyapatite crystals in bone tissue, which provide stiffness and compressive load resistance. In tendons and ligaments multiple fibrils make up collagen fibre, formed with the aid of proteoglycans.

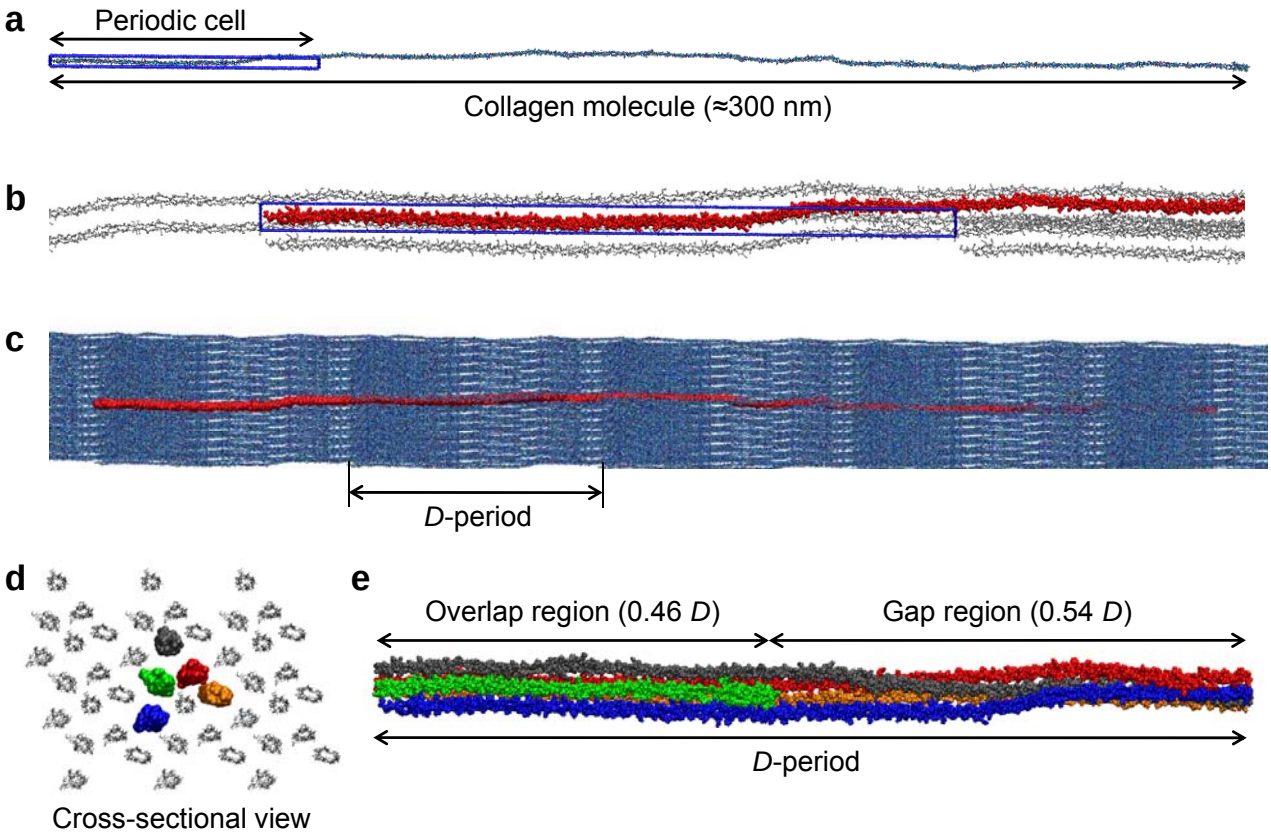


Figure 2 | Atomistic model of the collagen microfibril. The full-atomistic model of the collagen microfibril is generated starting from the *in situ* structure of the backbone geometry of full length collagen type I molecule as identified by x-ray diffraction, and using the associated information on the naturally occurring crystallographic unit cell ($a \approx 40.0$ Å, $b \approx 27.0$ Å, $c \approx 678$ Å, $\alpha \approx 89.2°$, $\beta \approx 94.6°$, $\gamma \approx 105.6°$) [6] (PDB ID 3HR2). **a**, Homology modeling is used to obtain the full-atom structure of the human collagen type I molecule. The collagen supramolecular model of the microfibril is generated by the periodic repetition of the unit cell. Panel **b** shows a portion of the original collagen molecule in red, while four periodic images of the molecule are represented in gray. The molecular packing topology obtained by the periodic repetition of the unit cell lead to the well known *D*-banding periodicity seen in AFM images of collagen microfibrils as shown in panel **c** (in red the original molecule, in blue the periodic images). **d**, Quasi-hexagonally packing of collagen molecules, which interdigitates with neighboring molecules to form a supertwisted right-handed microfibril as depicted in panel **e**. This image is obtained wrapping all collagen atoms (which spans several periodic units, see panel **a**) into a unit cell in order to visualize the microfibril periodic unit.

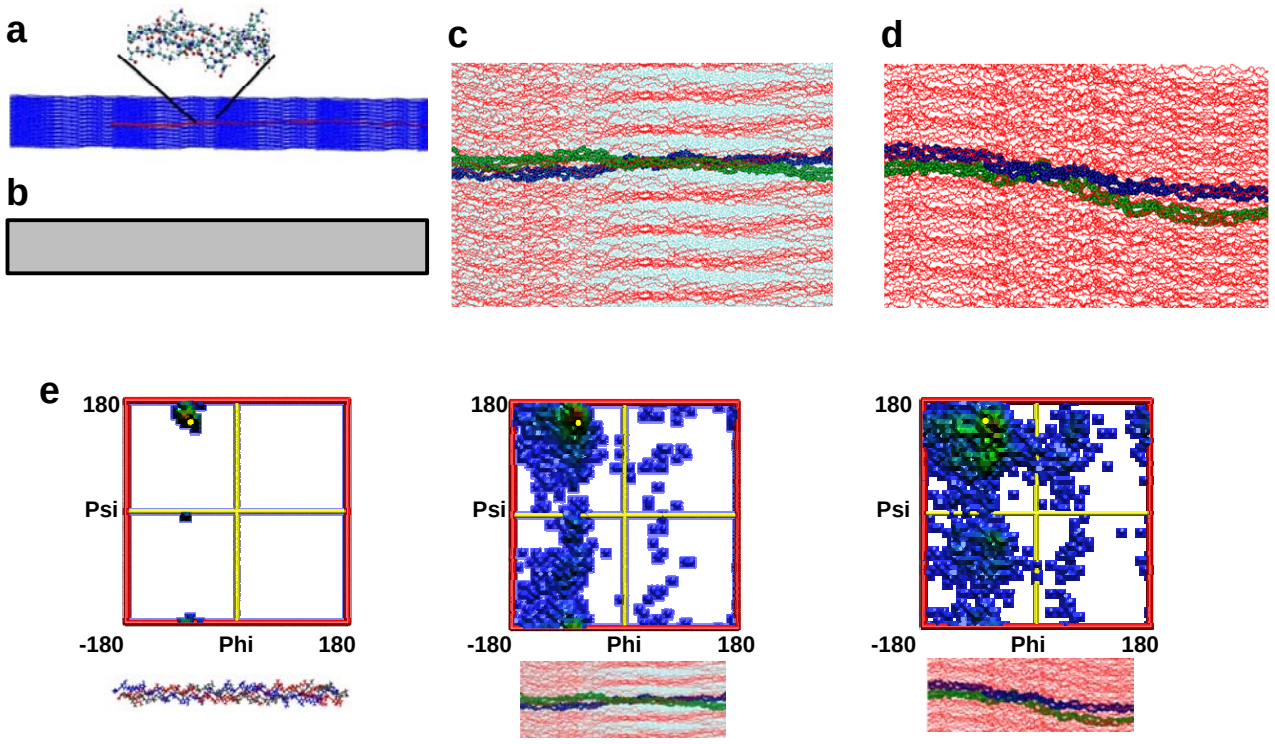


Figure 3 | Structural analysis and validation of atomistic collagen microfibril models. Comparison of the *D*-periodic banding observed for the full-atomistic microfibril model (panel **a**) and experimentally with SEM techniques (panel **b** removed here due to copyright restrictions). Panel **c** shows a detailed view of the equilibrated structure in proximity of the gap-overlap region, showing collagen molecules (in red, plus two highlighted molecules in blue and green) and water molecules (cyan). Panel **d** shows a snapshot of the gap-overlap transition region for the equilibrated dry collagen microfibril, which represents a much denser packing of molecules. Ramachandran diagram for a short collagen like peptide (left, panel **e**) and for the hydrated (wet) full atomistic microfibril (center, panel **e**) and for the dehydrated (dry) collagen microfibril (right, panel **e**), showing that the configuration is close to that of the polyproline II chain (Psi ≈150°, Phi ≈-75°, yellow dot) and thus close to the expected configuration of a collagen molecule. The dehydrated (dry) microfibril shows a more disperse distribution of dihedral angles.

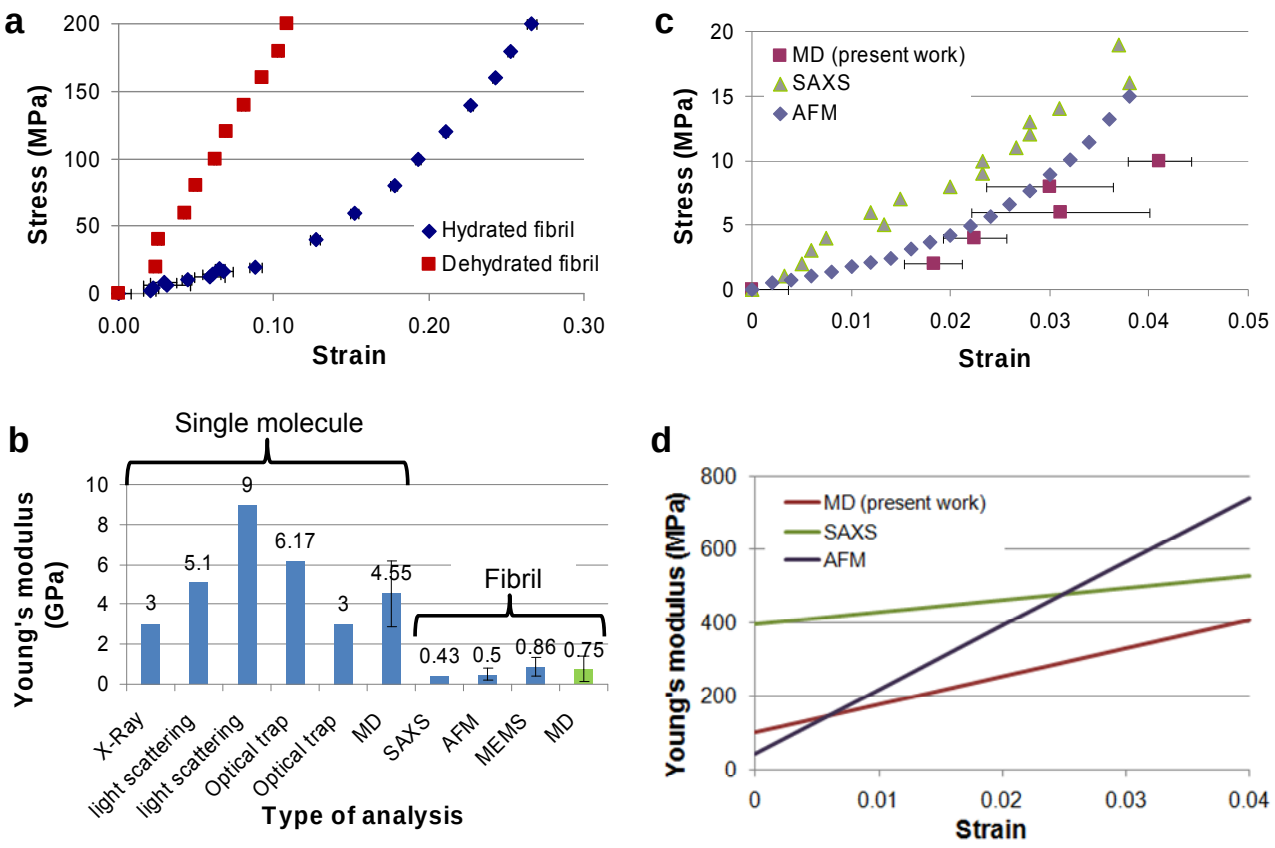


Figure 4 | Collagen microfibril stress-strain behavior, comparison with single molecule mechanics, and quantitative comparison with experimental results. The mechanical properties of both hydrated (wet) and dehydrated (dry) collagen microfibrils are determined imposing an increasing mechanical stress (negative pressure) along the fibril axis while maintaining the pressure on the other axes constant at 1 bar. Mechanical testing yields a fibrillar small-strain Young's modulus of ≈300 MPa and a large-strain modulus of ≈1.2 GPa for the hydrated (wet) model, whilst an almost linear behavior and an elastic modulus of ≈1.8 GPa (approaching 2 GPa for larger strains) is found for the dehydrated (dry) model (panel **a**). This finding suggests that the dry collagen microfibril tends to have a greater stiffness, a finding that is in agreement with experimental results (Table 1). **b**, Direct comparison of the Young's modulus obtained for solvated single molecules and microfibrils, featuring various experimental results and the predictions from our microfibril mechanics model. The calculated Young's modulus for the solvated collagen microfibril (green) results in very good agreement with experimental findings based on a variety of techniques including SAXS, AFM, and MEMS testing, which yield a small strain fibril Young's modulus in the range of few hundred MPa. **c**, Direct comparison of stress-strain curves obtained in this work and those obtained with experimental techniques. **d**, Young's modulus over strain (obtained from the gradient of stress-strain curves), comparing experiment and simulation for hydrated (wet) microfibrils.

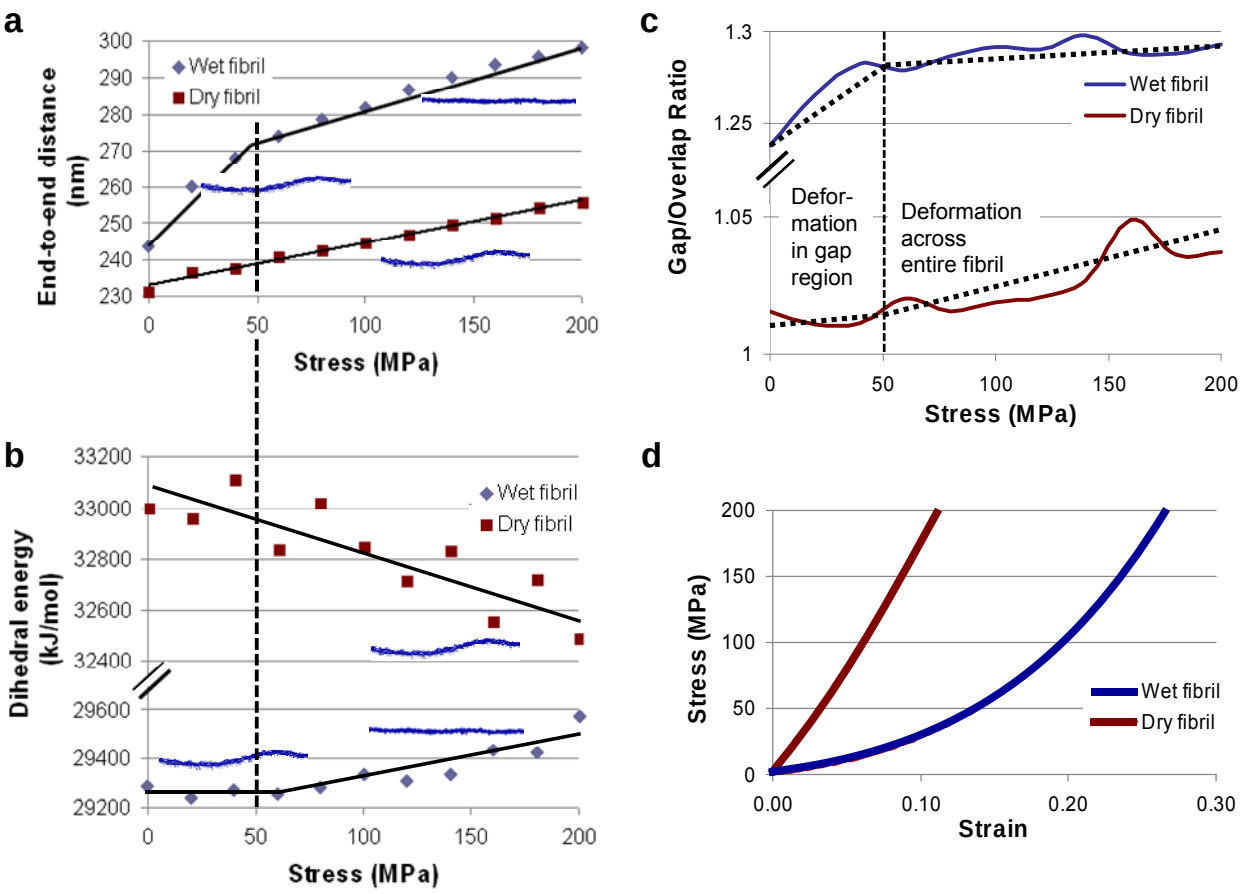


Figure 5 | Molecular deformation mechanisms during stretching for hydrated (wet) and dehydrated (dry) collagen microfibril. The two regimes observed in the hydrated (wet) microfibril stress-strain curve (showing a larger modulus for strains in excess of 10%, corresponding to an applied stress of ≈50 MPa) can be explained analyzing the behavior of single molecules during microfibril stretching. At small deformation the collagen molecule end-to-end distance increase linearly until the fibril stress reaches ≈50 MPa. Around this point the collagen reaches its contour length. Beyond this point, the molecular end-to-end still increases but the slope of curve is distinctly smaller (blue, panel **a**). This is due to the fact that below ≈10% strain the collagen molecule is straightened within the microfibril, while beyond this point the molecule is actually stretched, resulting in a larger resistance of the whole microfibril. The analysis of the dihedral energy of the systems shown in panel **b** confirm this observation, showing that for stress larger than 50 MPa the dihedral energy increase and thus that the molecule is deformed (blue, panel **b**).The molecular straightening is concentrated in the gap regions as shown by the increase in the gap/overlap ratio in the low-strain regime (blue, panel **c**). This suggests a microfibril deformation mechanism in which mechanical load initially straightens collagen molecules, particularly kinks formed in the gap regions, leading to an increase in the gap-to-overlap ratio. For larger loads, collagen triple helices undergo stretching resulting in larger microfibril stiffness (panel D). In the dry microfibril, the molecular end-to-end distance (red, panel **a**) increases linearly in the stress range analyzed, while the dihedral energy decreases (red, panel **b**). This suggests that in the dry microfibril the deformation mechanism initially involves primarily the straightening of the collagen molecules and not stretching of the molecules itself (this is confirmed by the observation that the end-to-end distance at 200 MPa stress is 260 nm, much shorter than the collagen molecules' contour length). The analysis of the gap/overlap ratio (red, panel **c**) shows that the deformation is initially distributed in both the gap and

overlap regions (since the ratio remains constant), and deformation affects the gap region only for larger stresses as shown by the increase in the gap-to-overlap ratio.

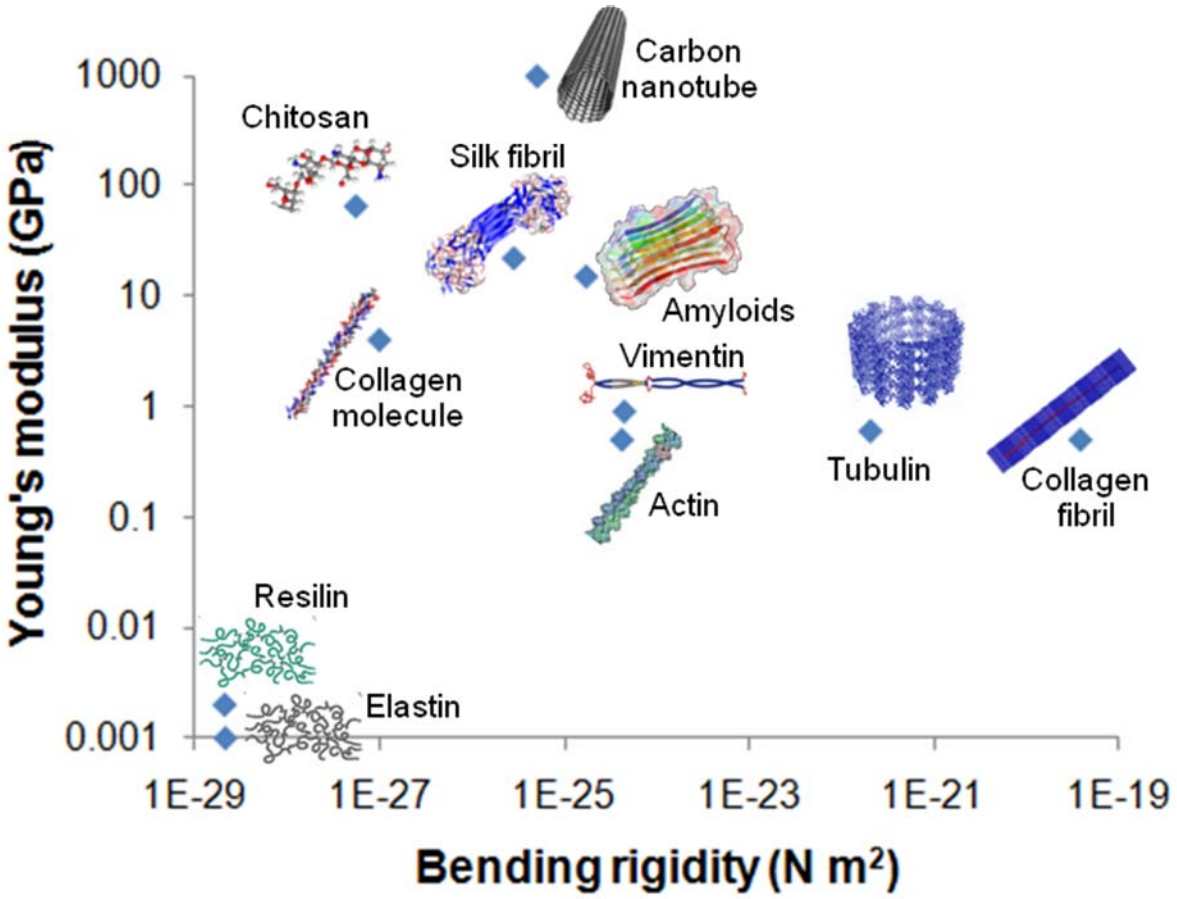


Figure 6 | Mechanical properties of materials at the nanoscale, comparing both biological and synthetic materials. Biological fibrils and fibres present a very vast range of mechanical properties in terms of Young's modulus and bending stiffness [71-75]. However, most of the protein materials feature Young's moduli in the range of 100 MPa to 10 GPa, well below the stiffness of many synthetic nanostructured material such as carbon nanotubes. On the other hand, bending stiffness shows a much greater variability. Collagen microfibrils and fibrils present a significantly enhanced bending rigidity at with only a relatively small decrease in the Young's modulus compared to a single molecule, showing a greatly effective fibril packing. The analysis shows that collagen fibrils provide a significant bending rigidity at relatively high Young's modulus.